THE UNIVERSITY OF ALBERTA

CLUSTERING IN FREE RECALL FOLLOWING PAIRED-ASSOCIATE LEARNING

by

© PEGGY ANNE RUNQUIST

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF PSYCHOLOGY

EDMONTON, ALBERTA

AUGUST, 1968

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "Clustering in Free Recall Following Paired-associate Learning" submitted by Peggy Anne Runquist in partial fulfillment of the requirements for the degree of Master of Science.

ABSTRACT

This experiment investigated the organization of stimulus and response terms into subsets during paired-associate learning on the basis of similarity relationships among the stimulus terms. The criterion for organization was the occurrence of clustering in free recall of the stimulus and response terms following paired-associate learning.

Six groups of 20 Ss learned 16-item paired-associate lists to a criterion of 14/16 correct. After a 3-min. interval filled with arithmetic problems, Ss were asked to recall either the stimulus terms or the response terms. A second test for the other members of the pairs followed immediately. This recall procedure was repeated on a third and fourth recall trial.

For two groups, the stimuli were unrelated nouns. For four groups, the stimuli were four instances of four mutually exclusive conceptual categories. Two of these categorized groups had a series of learning trials on the categorized stimulus terms prior to paired-associate learning.

Two additional groups received pretraining trials only, on either just the stimulus terms or just the

response terms, followed immediately by the arithmetic problems and the free recall tests.

The groups that learned the categorized lists recalled more stimulus and response terms on the free recall trials than the groups that learned the noncategorized lists. Stimulus pretraining increased the number of stimulus terms recalled, but did not affect response recall.

The categorized paired-associate groups showed significant clustering of both the stimulus terms and the response terms when measured in relation to the stimulus categories. Since the noncategorized paired-associate groups and the response pretraining group did not cluster the responses in the same way, the clustering of the response terms by the categorized groups can be attributed to organization of the responses into subsets based on characteristics of the stimuli with which they were paired.

ACKNOWLEDGEMENTS

I would like to thank Dr. W. W. Rozeboom for his advice and critical comments on the writing of this thesis.

The research was supported in part by Research grant APB-88 from the National Research Council of Canada to Dr. W. N. Runquist.

TABLE OF CONTENTS

CHAPTER PAGE

LIST OF TABLES

Chapter I

Introduction

The purpose of this thesis research was to investigate the existence of organizational behaviour in paired-associate learning. More specifically, the problem was to determine if the stimulus and response terms would be grouped into subsets based upon specifiable stimulus characteristics during paired-associate learning.

Although psychologists have long been concerned about how organization affects memory, it has only been quite recently that any major attempts have been made to determine the nature of this organization or to quantify its effects. In fact, it was not until Miller's work on information processing that the need for recognizing some kind of organizational process in memory was generally appreciated. Miller (1956) demonstrated that there is a limit to the number of discrete items of information that any individual can handle at any one time; this limit he estimated to be 7 ± 2 chunks. Obviously then, to handle more information, the individual must recode or reorganize the input in some way so that the amount of information in each chunk is increased. Memory would then consist of the recall of these 7 ± 2 chunks and the retrieval of their

Chapter 1

Introduction

The purpose of this [illegible] research was to demonstrate the existence of organizational processes in paired-associate learning. More specifically, the problem was to determine if the stimulus and response terms would be encoded into [illegible] [illegible] [illegible] [illegible] during paired-associate learning.

Although psychologists have long been concerned with how organization affects memory, it has only been quite recently that any major attempts have been made to understand the nature of this organization or to quantify its effects. In fact, it was not until Miller's work on information processing that the need for recognizing some kind of organizational process in memory was generally appreciated. Miller (1956) demonstrated that there is a limit to the number of discrete items of information that any individual can handle at any one time; this limit he estimated to be 7 ± 2 chunks. Therefore, then, to handle more information, the individual must recode or reorganize the input in some way so that the amount of information is then within his limitations. Memory would then consist of the recall of some 7 ± 2 chunks and the retrieval of their

contents. As Kendler (1966) has pointed out, there is no longer any disagreement about whether memory is organized; the problem now is what kind of mechanism brings this organization about.

One of the first questions to consider is what is meant by the term organization. Mandler (1967) has used the term very generally: "A set of objects or events are said to be organized when a consistent relation among the members of the set can be specified and, specifically, when membership of the objects or events in subsets (groups, concepts, categories, chunks) is stable and identifiable (p. 330)" The major problem with this description when applied to a verbal learning situation is that it does not specify whether "organization" refers to some property of the stimulus material itself, or whether it refers to some activity engaged in by S. Clearly, it is this latter sense of the term, i.e., organization as a response, that is of major interest in this kind of research. "Organization" in the former sense, however, is not in the least ignored. Properties of stimulus materials, or more precisely, relationships among the stimulus items, are thought to be the major determiner of the kinds of organizational activities that are engaged in,

and, as such, could be referred to as the "bases" of organization. The present study, in fact, deals with only a limited domain of organizational behaviours, i.e., those for which the probable basis of organization can be specified.

While there have been several approaches to the study of organization and memory, one of the most productive has been the study of clustering in free recall. In a typical free recall learning situation, the _S_ is presented with a list of words, one at a time, and then asked to recall as many words as possible, and in any order. This procedure offers two advantages for the study of organizational behaviour: first, the opportunity for any organizational activities on the part of _S_ to be shown in his recall performance is much greater than in the other, more structured, learning tasks; and secondly, the effects of at least some organizing activities can be measured by examining the order in which he recalls the words (Tulving, 1968). The assumption made in free recall research is that recall order reflects the order in which the _S_ thinks of the words. Clustering has usually been defined as the occurrence in recall of sequences of two or more items which are somehow related to one another, and

it is presumed to represent the S's tendency to reorganize his recall on the basis of pre-experimental associations or conceptual relationships. Clustering has been repeatedly demonstrated over a wide range of stimulus materials and variations in procedures, and has been consistently accompanied with an increase in the number of items that are recalled (e.g., Mandler, 1967).

Paired-associate learning, which is usually conceptualized as the acquisition of several relatively isolated associations, has also shown some evidence of organizational processes under certain conditions. Recent studies (e.g., Runquist, 1966; Brown & Battig, 1966) have shown that with paired-associate lists composed of several subsets of items, with items within a given subset being related to one another and items from within different subsets being unrelated to one another, both acquisition performance and long-term retention were facilitated under conditions that were expected to produce greater interference. These results were interpreted as being evidence for organizational behaviour or "grouping" in paired-associate learning, with the basis of this grouping being the relationships among the items within a given subset. These studies

have reportedly shown "grouping" on the basis of differential degree of learning, differential frequency of appearance, and formal and meaningful similarity. The evidence that this facilitation is actually a result of some type of organizational process, however, has been indirect and fragmentary, consisting primarily of a high proportion of confusion errors within a subset, or some extremely complex transfer phenomena.

In the present experiment, an attempt was made to obtain somewhat more direct evidence for the organization of stimulus and response terms into subsets by means of the incorporation of a free recall test following paired-associate learning. The rationale was simply that, if the stimulus and response terms are indeed organized during paired-associate learning, this organization ought to be manifested as clustering on the free recall test.

Chapter II

Organization in Free Recall

Three types of paradigms, differing primarily in the class of materials used to induce clustering and in the method of evaluating this clustering, have been used in most of the studies concerned with organization in free recall. Specifically, these are: categorical clustering, associative clustering, and free organization. This chapter will outline each of these procedures, summarize some of the kinds of variables that have been found to influence the amount of clustering, and conclude with a brief presentation of the theoretical interpretations that have been offered in explanation of these clustering phenomena.

Categorical Clustering

Categorical clustering was first observed by Bousfield and Sedgewick (1944) in a study concerned with the characteristics of sequences of associative responses. They found that the Ss, when asked to list items in specified categories, tended to respond in sequences of subcategories. For example, if instructed to list a series of animals, the S would first respond with several canines, then several felines, etc. In 1953, Bousfield set up a research program designed specifically

to develop techniques for inducing clustering and measures to quantify the amount of clustering that occurred. In his first experiment, (Bousfield, 1953) the Ss were read aloud a list of 60 nouns composed of 15 examples of each of four mutually exclusive categories (animals, names, professions, and vegetables) presented in a random order. The Ss were then given a single recall trial on which they were to write down as many words as they could remember. To assess the amount of clustering Bousfield developed the "repetition" measure. A "repetition" refers to the contiguous recall of two items from the same category. The amount of clustering for each S is determined, then, by summing, over all the categories in the list, the number of "repetitions" that occur.

The categorical clustering procedure, then, is characterized by the type of stimulus list used, this list predominately being composed of concrete nouns representing two or more mutually exclusive conceptual categories (e.g., animals, vegetables, etc.). Bases of categorization other than conceptual category membership have also been used to obtain clustering and have met with varying degrees of success. Some of the relationships among items which have produced clustering are synonyms

(Cofer, 1959); response dominance (Bousfield and Puff, 1964); and structural characteristics of geometric designs (Bousfield, Berkowitz, and Whitmarsh, 1959). Form class or parts of speech as a basis for organization, on the other hand, has not produced clustering (except under certain conditions which will be described later), presumably because it is too broad a category for the relationship among items to be perceived (Cofer and Bruce, 1965).

In the following section, several variables which have been found to influence the amount of clustering will be briefly discussed. In the order in which they will be presented, these variables are: normative frequency, exhaustive vs. nonexhaustive categories, number of categories, blocked vs. random presentation, multiple-presentation of list, presentation rate, interval between study and test trial, and the effect of context.

Normative frequency. Cohen, Bousfield, and Whitmarsh (1957) obtained cultural norms from 400 undergraduates at the University of Connecticut who responded to 43 different category names with the first four specific associates to each that they could think of. These responses were then ranked on the basis of

their frequency of occurrence to each of the category names. Several studies (Bousfield, Cohen, and Whitmarsh, 1958; Bousfield, Steward, and Cowan, 1964; Cofer, Bruce, and Reicher, 1966) have demonstrated both superior recall and clustering in lists comprising high frequency responses from these norms than for lists of low frequency responses.

Exhaustive vs. nonexhaustive categories. Exhaustive categories are those whose instances exhaust all items generally subsumed under the category label, e.g., north, south, east, west. Cohen (1963a, 1963b) has shown that, although the mean number of stimulus categories that are represented in recall by at least one item from each category was the same for both types, significantly more words within a category were recalled when exhaustive categories rather than nonexhaustive categories were used.

Number of categories. The relationship of the number of categories in the list to recall and amount of clustering is not entirely clear since it varies with list length, whether recall is cued (i.e., category name given at recall) or not, and with prior experience of Ss in clustering experiments. Generally, however, there is an inverse relationship between number of

categories and mean recall with long lists (24 or more items) and non-cued recall, when naive Ss are used (Dallett, 1964, Exp. IV & V; Bousfield and Cohen, 1956). With cued recall and short lists (i.e., 12 items), mean recall is directly related to the number of categories (Matthews, 1954). And with noncued recall and short lists, the relationship is curvilinear, with the best performance with two to four categories (Dallett, 1964, Exp. I-III).

Blocked vs. random presentation. Blocked presentation refers to the experimental situation in which all members of a given category are presented contiguously in the stimulus list. The order of categories and the order of items within each category can, however, be randomly varied. Three studies (Cofer and Bruce, 1965; Cofer, Bruce, and Reicher, 1966; and Dallett, 1964) have found both recall and clustering to be superior with blocked as opposed to random presentation. Using parts of speech as the basis of categorization, Cofer and Bruce (1965) showed less than chance level clustering under random presentation and only minimal clustering under blocked presentation. Over several trials, however, the difference in the amount of clustering produced by these two methods was significant.

<u>Multiple presentation of the list</u>. There are two ways in which number of stimulus list presentations has been varied in studies of clustering. The first method involves several presentations prior to a single free recall trial, and the second involves a series of alternating study-recall trials. Two studies have used the first method. Bousfield and Cohen (1955) presented a four-category, 40-item list either 1, 2, 3, 4, or 5 times prior to a single 10 min. recall period. Both mean number of words recalled and mean amount of clustering were found to increase directly with number of presentations. Using synonyms as the basis for inducing clustering, Cofer (1959) presented the list either once, twice, in the same order, or twice but in different orders before recall. The results showed higher recall and clustering for <u>S</u>s who saw the list twice, and this difference was even greater when presentation order was the same on the two study trials. Using the alternating study-test method, Bousfield, Berkowitz, and Whitmarsh (1959) showed that both clustering beyond a chance level and mean recall increased progressively as a function of trials.

<u>Presentation rate</u>. The Cofer, Bruce, and Reicher (1966) study is the only one in which clustering as a

function of different presentation rates has been investigated. Using rates of 1, 2, 4, and 4.4 sec., they found both mean recall and clustering increased with slower rates, although this relationship varied somewhat with method of presentation (i.e., blocked vs. random) and with normative frequency.

Interval between study and test trial. Gonzales and Cofer (1959) found that if a second recall test was given 5 min. after the first test, clustering, but not recall, tended to increase. Cofer, Bruce, and Reicher (1966) also found clustering on a second trial to increase, and showed that this clustering was significantly greater than that of a control group which waited an equivalent length of time but did not have the interpolated recall test.

Two studies have compared the retention of categorized lists after longer intervals. Cofer (1959) found a marked decrease in clustering after a two week interval; however, both the number recalled and the amount of clustering were still greater for a group learning a categorized list than for a group learning a noncategroized list. Brand and Woods (1959) compared the retention of a categorized list at 1, 2, and 3 week intervals. One group was tested at just one of the intervals. The repeated testing group showed an initial

decrease in clustering at the one week interval, followed by a progressive increase on the second and third weeks. For this group, there was a drop in recall at the first retention interval, but no further change after that. The independent groups showed a progressive decrease in both clustering and recall as the interval between study and test trial increased.

The effect of context. Gonzales and Cofer (1959) conducted a series of experiments to study the effects of context upon clustering. Starting with a basic 40-item list composed of nouns in four categories, they modified the procedure somewhat by presenting adjective-noun pairs rather than just the nouns alone, and by asking for recall of either member of the pair. The results can be summarized in terms of the effects of five different kinds of pairings. (i) Specificity effect. When all the adjectives were chosen so as to commonly modify only one of the nouns in the list, and the adjectives were unrelated to each other, clustering of the nouns was reduced to chance level. (ii) Mediation effects. When a list of uncategorized nouns, which did not cluster when presented alone, were modified by adjectives representing discrete categories, which did cluster when presented alone, the recall of

the nouns also revealed significant clustering when scored on the basis of the modifying adjectives. The reverse was also true: uncategorized adjectives would cluster on the basis of the nouns they modified. These effects were independent of word order, i.e., noun-adjective or adjective-noun. (iii) Mediated facilitation effects. When clustering adjectives and clustering nouns were presented so that the categories were compatible with each other, clustering and the recall of nouns was facilitated. These effects did not occur with adverb-verb pairs nor with four word combinations, e.g., adjective-noun-verb-adverb. (iv) Mediated conflict effects. When clustering nouns and clustering adjectives were incompatible with each other, i.e., adjectives from a given category modified nouns from all four categories, clustering and recall were impaired. (v) Inappropriate modification effects. When the adjective modifying a given noun was inappropriate, e.g., leafy dog, both clustering and recall were impaired. Later association data collected by Cofer (1960) showed that, with the exception of the mediated facilitation effect, the results of this study can be accounted for fairly well on the basis of changes in the associative relationships involved by

the modifying adjectives.

Associative Clustering

Jenkins and Russell (1952) were the first to investigate the influence of associative strength as a factor in clustering. They presented twenty-four stimulus-response pairs of words from the Kent-Rosanoff word list in a restricted random order (the restriction being that a response word could not follow its stimulus). Clustering was said to occur when the two words in a given pair occurred together at recall. A chance baseline, against which the occurrence of both forward and backward associations was compared, was determined by the number of occurrences of arbitrary pairs, i.e., a stimulus word followed by a specified response, other than its own, selected randomly from the list. The results showed that the recall of the associated pairs was significantly higher than that of the control pairs, and that the number of forward sequences recalled was significantly greater than the number of backward sequences.

Associative clustering, then, is also characterized by the type of stimulus list used; in this case, however, the list is composed of pairs of associatively related words (as determined by association norms) which do not

necessarily belong to the same conceptual category.

The occurrence of clustering for pairs of associatively related words has been demonstrated in both adults (Jenkins and Russell, 1952) and children (Wicklund, Palermo, and Jenkins, 1965). Moreover, clustering can also occur when the two words of a pair are related, not directly, but through a third word not appearing in the list (Cramer, 1965). Bousfield, Whitmarsh, and Berkowitz (1960) have also shown that associative clustering can be predicted on the basis of the extent to which two words elicit common associates. Bousfield, Steward, and Cowan (1964) have shown this measure to be a better predictor of categorical clustering than an index of inter-item association, i.e., the extent to which the items in the list elicit one another.

Associative clustering has not been extensively studied, and only a few influences have been determined:

<u>Associative strength</u>. The Jenkins, Mink, and Russell (1958) and Wicklund, Palermo, and Jenkins (1965) studies have both demonstrated that the occurrence of clustering in both forward and backward directions is a positive function of the associative strength between the two words of a pair. Even clustering of the low

associative strength pairs was above chance level for adults (Jenkins et al, 1958) but it was only at a chance level for children (Wicklund et al, 1965).

Multiple presentation of the stimulus list. Rosenberg (1966) investigated associative clustering in a 48-item list over five alternating study-test trials. Clustering was found to increase as a function of trials, although this relationship did vary somewhat as a function of sex, direction of association, and associative strength.

The effects of set. Marshall and Cofer (1961) organized lists according to different levels of the mutual relatedness index, MR (the number of assocations two words have in common in proportion to the total number of associations). When Ss were given a set to organize their recall on the basis of relations in the list, clustering was found to increase at high MR levels but not at low MR.

Free Categorization

Mandler (1967) developed the free categorization technique in reaction to what he felt was a serious limitation of procedures which investigate the occurrence of only experimenter-defined organization, i.e., the problems of interpreting negative findings. The

question that Mandler raised was, if the S does not cluster, does this finding mean that he has not discovered the E-defined categories, or simply that he chooses not to use them because he has found some other basis of organization which is more useful to him. If the latter alternative is true, clustering measured in terms of E-defined categories may underestimate the actual degree of organization that occurs.

For this reason, Mandler attempted to find a measure for directly studying the organization imposed by Ss. He reports a series of experiments, all of which follow the same general method as the Mandler and Pearlstone (1966) study. In these experiments, Ss were first asked to sort a list of words into a number of categories. Usually the Ss were free to use from two to seven categories and any system of sorting they wished. The sorting task continued until S made exactly the same sort on two consecutive attempts. After reaching criterion, Ss were asked to recall all of the words in the list in any order they wished. Clustering was evaluated in terms of the S-defined categories.

The results of Mandler's set of studies can best be summarized as follows: (i) When the S was

unconstrained as to the number of categories, the mean number of categories used in the sorting task was about 4.5. (ii) The amount recalled increased as the number of categories increased from 2 to 7, and there was a tendency for recall to decrease when more than seven categories were used. (iii) Recall was independent of the number of trials to criterion on the sorting task. This finding plus the second would argue that recall is a function of the kind of organization imposed by the learner and not the number of trials taken to attain that organization. (iv) There was significant clustering on the recall trial when clustering was evaluated in terms of the categories used in the sorting task.

In a recent experiment (Mandler, 1967), the effects of instructions to categorize and/or recall were determined. Ss were instructed to write the words on a sheet of paper divided into seven columns as the words were presented. Four groups were run: Category-Recall; Category-No Recall; No Category-Recall; and No Category-No Recall. The Category group was told to categorize the words as they were presented by writing words that went together in the same column. The No Category group was told to write the first word in the

first column, the second in the second column, etc. The Recall group was told that recall would be tested later, and the No Recall group was not warned of the recall test. The major finding was that the No Category-No Recall group recalled fewer words than the other three groups, which did not differ from each other. In the two Category groups, the amount recalled was again a function of the number of categories used by S. Mandler interpreted these results as showing that "both recall or organizing instructions produce equivalent organization and equivalent recall (p. 357)."

Theoretical Interpretations

There have been three major interpretations of clustering phenomena: perception of the superordinate, coding interpretations, and associative interpretations. The phenomenon that they seek to explain is why clustering occurs and how it is able to affect recall.

The superordination interpretation was primarily used in the early Bousfield papers (Bousfield, 1953; Bousfield and Cohen, 1953). The clustering phenomenon was discussed in terms of two variables: (a) the habit strength between the item and the category to which it belongs, this habit strength being a function of the number of reinforcements this item has received

both before and during the experiment, and (b) the relatedness increment or generalized habit strength between that item and other similar items. The strength of these two variables determines the probability of whether the next item recalled will be from the same or a different category, thus determining the amount of clustering obtained. Phrased in terms of a categorical clustering experiment, the recall of a single word tends to activate a superordinate, i.e., the category represented by the words which belong to it. Once this superordinate is activated, it will tend to facilitate the perception and recall of the words which belong to the same category by virtue of the generalized habit strength between similar items.

Somewhat related to this position are the coding interpretations of clustering of Cohen (1966) and Mandler (1967), both based on Miller's (1956) concept of chunking. Cohen (1966) has pointed out that while the number of recalled words greatly exceeds the so-called span of immediate memory, the number of categories represented by the recalled words is relatively fixed. Such results according to Cohen, suggest that Ss readily detect the categorized nature of the list, store some coded representation of each category in memory, and

attempt to retrieve this coded information during recall. Mandler's position is similar. He agrees with Miller that memory is arranged hierarchically, and that the storage capacity within any one level is limited, but argues that there are more than just two levels of organization. In addition to chunks and items within chunks, the chunks may themselves be organized into "superchunks" and so on. There is a limit to the number of levels of organization within any one "schema", and this limit is postulated as being equal to the limit in the number of items within any one level. The "magical number" here is 5 ± 2. The above reasoning implies that if more than five words are recalled, some of the words were organized into chunks; if more than 25 words are recalled, the chunks themselves are organized into superchunks, and the ultimate limit within any one "schema" is about 5^5 or 3125 words. Another assumption made by Mandler is that chunking proceeds primarily by way of categorization of sets of words. That is, when the S is faced with the task of memorizing a list of words, he organizes them into a set of conceptual categories. A category, then, is taken as the equivalent of one of Miller's chunks.

While it has already been noted that associative factors can influence clustering, the question may be raised as to whether clustering can be explained solely in terms of associative relationships or whether it is necessary to include an additional concept such as superordination or coding. Two studies have been addressed to this problem. Both studies used pairs of words matched either for associative strength (Bousfield and Puff, 1965) or associative overlap (Marshall, 1963). For half of the pairs, the words belonged to the same taxonomic category, while for the other half, the words were not from the same category. Both experiments showed that the categorized pairs clustered to a greater extent than the non-categorized pairs even though the associative factors had been equated. The investigators concluded that, at least in some siutations, some concept such as superordination or coding is necessary to explain the data. Cofer (1966), however, suggests that superordination can itself be explained on the basis of associative relationships. His position is that, in addition to interword associations existing between the instances of a given category, there are also associations between an instance of a category and the

category name (or some representation of this name). According to this view, then, all associative factors were not equated in the studies just described; i.e., the only associations taken into account were those among the stimulus words themselves and not the associations between the word and the category to which it belonged.

Another question that can be raised in this context is whether the recoding interpretations of Cohen and Mandler are incompatible with associative explanations. The basic notions of these interpretations are that: 1) the material to be remembered is "recoded" into chunks or categories; 2) that it is these chunks that are "stored" in memory and subsequently recalled; and 3) clustering represents the "retrieval" of the information stored in successive chunks. Obviously, associative relationships among items could be a major determiner of which items are stored together in a given chunk, and the primary means by which the contents of the chunks are retrieved. To the extent that the relationships among stimulus items determine how the items are organized, then, these interpretations are not in the least incompatible with each other.

Summary

This chapter was concerned primarily with a description of the three basic procedures used to induce clustering in free recall and some of the variables that have been found to influence this clustering.

Regardless of differences among these procedures, all three have demonstrated organizational behaviour in the free recall learning situation and have attributed the bases of this organization to relationships among the stimulus items. One question that is not readily answered at this point is whether some special organizational process must be incorporated to explain this organizational behaviour or whether associative principles alone will suffice.

Chapter III

Evidence for "Organization" in PA Learning

In the typical paired-associate experiment, the S is shown several pairs of verbal units, one at a time. The left-hand member of the pairs is designated as the stimulus term and right-hand member as the response term. The task of S is to learn to associate the two members of the pairs so that, when the stimulus term is presented alone, he will be able to recall the response term that was paired with it.

Although, the paired-associate task is often conceptualized as a "straightforward extension of simpler S-R principles (Battig, 1968)", it is generally accepted now that there is much more involved than the simple formation of stimulus-response associations within individual pairs of items. A more recent trend, in fact, has been toward a search for and the identification of the component processes involved in acquisition and recall (Postman, 1968). In paired-associate learning, several subprocesses have been postulated (e.g., Battig, 1968), but only three of these have gained any wide-spread acceptance, i.e., response learning, associative learning, and stimulus differentiation. Recently, however, there has been

accumulating a great deal of evidence for the existence of yet another process involved in paired-associate learning, namely, that of interpair grouping. Described by Battig (1968), interpair grouping refers to the development of a "hierarchical organizational structure, whereby the overall list is subcategorized into subsets or groups of pairs with some common property or interrelationships among them (p. 63)."

It has been well-known for some time that paired-associate learning is affected by similarity and other interrelationships between items. The usual conception, however, is that such interrelationships are a major source of interference in both learning and retention. According to this view, the effects of these relationships get eliminated during the course of learning, and, at the end of learning, each pair represents a single associative unit which is independent from all the other pairs in the list.

The notion of an organizational process, such as interpair grouping, that is developed or strengthened rather than being eliminated is definitely in contrast to the foregoing position. Yet, recently, there has been increasing evidence that not only does such a process exist, but also that its effect is to

facilitate both acquisition and retention performance on paired-associate lists. This facilitation is presumed to be due to a reduction in the amount of interference, brought about by grouping the items into categories in such a way that most of the interference comes from the restricted set of items within each category and very little from items outside the category (Runquist, in press (a)).

The evidence for this organizational process is, for the most part, both fragmentary and indirect, coming largely in the form of occasional facilitation with categorized or groupable lists and an increasing proportion of the total overt errors which are generalized errors with these lists. As used in the following discussion, a generalized error is said to occur when the S responds to a given stimulus item with the response item that is appropriate to another item belonging to the same stimulus category. The greater part of this chapter will briefly describe those experiments which have found evidence for the process of interpair grouping. It should be stressed at this point that these studies have been able to demonstrate grouping behaviour because the items to be learned were in some way groupable. Accordingly, in the following

discussion, emphasis will be placed on the particular types of evidence found, the bases on which the items were grouped, and the interpretations of how grouping would affect learning.

Grouping on the Basis of Similarity Relationships

Underwood, Ekstrand, and Keppel (1965, Exp. V) found evidence for grouping in the form of within-category or generalized errors while investigating the effects of conceptual similarity in learning. In this experiment, 12-item lists comprised of 1, 2, 3, 6, or 12 different concepts as the stimulus terms and double letters as response terms were used. The main result, so far as the investigators were concerned, was that as the number of concepts increased (or as number of stimuli within a category decreased), the rate of learning increased. More inportant for the present discussion, however, was the finding that the frequency of occurrence of within-concept errors was significantly higher than that which would be expected if S had responded randomly. These investigators interpreted this result as evidence for "S-R limitation" (or interpair grouping) on the basis of category membership, the effect of this process being to limit the number of possible responses which can be correct for a given stimulus.

The following four experiments, recently conducted at the University of Alberta, have shown evidence for grouping in paired-associate learning on the basis of formal similarity among the stimulus terms. Since the high similarity lists were constructed in the same way for all these experiments, it would be advantageous to briefly describe these lists before presenting the studies in any detail. In high similarity lists, then, the stimulus terms consisted of several subsets of two similar items (e.g., XUL and XIL; QET and QOT), in such a way that similarity within a stimulus subset was high and that between subsets was low. The two response terms which were paired with the similar stimuli were always dissimilar to each other. Lists of this type are designated as pairwise similar lists in the following studies.

Runquist (1966) reports the results of two experiments in which list length and pairwise similarity were varied in a factorial design. In the first study, the stimulus terms were low association value trigrams and the responses unrelated nouns. Subjects learned high or low similarity lists consisting of either 8 or 16 items. The results showed a significant difference favoring the low similarity group with the short

list, but no effect of similarity with the longer lists. Runquist interpreted this finding on the basis of grouping. In addition to the response learning, discrimination, and associative stages, he suggested that for the high similarity lists, there was also a stage (following response learning) in which the S learns which two responses go with each pair of stimuli (e.g., S may learn that "heaven" and "dinner" go with XUL and XIL, but not know which is which). Facilitation for this group could be expected for two reasons: by grouping the responses on the basis of the similar stimuli with which they are paired, interference from other sources in the list would be reduced; and, secondly, by knowing which response is correct for one of the stimuli, S would then know, by elimination, that the other response must be correct for the second stimulus. Added support for this interpretation was gained through examination of the overt error data: after the response had once been given correctly, over half the overt errors were generalized errors, i.e., the response to the similar stimulus. In the second experiment the need for learning which two responses went with each stimulus pair was eliminated by providing them to the S.

The S's task in this experiment was to indicate, by pushing either of two buttons, to which of two categories each stimulus belonged. With the response grouping under high similarity conditions already given to the S, there was some evidence that the high similarity group learned even faster than the low similarity group with longer lists. One question that could be raised about both of these studies, is why was the learning of the long lists facilitated and not that of the short lists. One plausible explanation would be that with short lists the individual associations are learned too rapidly for grouping to occur.

A more recent study (Joinson and Runquist, in press) compared the retention of lists varying in pairwise intralist stimulus similarity after one week, following learning to each of three performance level criteria. Similarity was defined in terms of identical letters in low association trigrams, and mean terminal performance level was exactly equated at the three levels for the low and high similarity conditions. The major finding of this study was that retention was substantially higher for the high similarity group at the medium criterion of learning, but not significantly different

at the low or high criteria. Significant, too, is the fact that this medium criterion group also had the largest number of generalization errors at recall. These data were interpreted to be the result of response grouping on the basis of pairwise similar stimuli. The reason given to explain why grouping might be expected to facilitate long term retention was that it would enable the S to recall which two responses were paired with the similar stimuli even when the particular associations have been forgotten. In the low similarity groups, on the other hand, this focussing would not be available. Furthermore, at low criteria, it was proposed the high similarity list had not had sufficient practice to be grouped, while at the high criterion, there was so little associative forgetting that the phenomenon had only a slight effect. Whether being able to recall group membership facilitated recall by reducing interference from other pairs in the list or simply by producing more efficient guessing could not be ascertained from the data.

One study (Runquist, in press (b)) has obtained more direct evidence for this grouping interpretation. In this experiment, Ss first learned one paired-associate list in which the stimulus term again consisted of

pairwise similar trigrams and the response were nouns, and then were transferred to another list in which the stimulus and response terms were re-paired. This re-pairing was done in two ways: a response term was re-paired with the other stimulus in the same set of similar terms (reversal re-pairing), or it was re-paired with a stimulus from within a different set (non-reversal re-pairing). The rationale behind this experiment was that if grouping occurs, one would expect reversal re-pairing, where a response term belongs with the same group of stimuli, to be easier than non-reversal re-pairing, where a response term is shifted to a new group of stimuli. The results supported this hypothesis: items re-paired with similar stimuli were learned much faster than those re-paired with dissimilar stimuli. If the competing response tendency had been extinguished in first-list learning, rather than being maintained, this facilitation in second-list learning, for the reversal group should not have occurred.

One of the questions that could be raised in connection with the preceding four studies is whether this grouping behaviour could be explained simply on the basis of stimulus generalization principles. This

interpretation would account for the high proportion of generalized errors during acquisition in terms of competing response tendencies resulting from generalization between the similar stimuli. According to this notion, however, these response tendencies are supposed to interfere with the learning of correct associations, not facilitate it. The usual conception, in fact, is that these error tendencies must be "extinguished" or "unlearned" in order for the correct associations to be acquired. The grouping interpretation, on the other hand, is that the response terms paired with the similar stimuli become directly associated or "grouped" together as a result of their association with the similar stimulus terms. In the re-pairing study, then, both the stimulus "groups" and the response "groups" remain intact and only the specific stimulus-response pairings are changed.

Runquist (in press (a)) has also conducted three other experiments which suggest that Ss respond to the "structure" of similarity relationships in paired-associate lists. In these studies similarity of stimuli (low association value trigrams) was defined on the basis of letter-position identity "rules", i.e., the number and position of identical letters among all the

pairs of stimulus terms in the list. In a previous study, Runquist and Joinson (in press) had obtained similarity ratings for pairs of trigrams representing different letter-position rules (e.g., a pair of trigrams with the rule "first two letters identical" received a rating of 66.3 out of a possible 100). The present experiments involved the use of several lists representing variations in total stimulus similarity (i.e., the sum of the similarity ratings between all the pairs of stimuli in the list), in the variability of the similarity ratings among the stimulus pairs within the list, and the number of "rules" contained in the list. The response terms in all cases were single digit numbers. In the first experiment, stimulus similarity was varied in single "rule" or homogeneous lists (e.g., all stimulus terms started with the same first letter). In the second experiment, lists were of moderate similarity, but varied in the number of rules and the variability of similarity ratings among the stimuli. In the third experiment, all lists contained four rules, but varied in similarity. In general, similarity was found to produce some interference in homogeneous (one-rule) lists. With multiple-ruled lists, the amount of interference was increased,

and in lists in which there was high variability of the similarity ratings and a constant number of rules, the amount of interference was reduced. These findings demonstrate quite definitely that variations in the structure of relationships among the stimulus items in the list do affect learning.

Variable Frequency

Two paired-associate experiments conducted by Runquist (1965, 1967) have shown that differential frequency of presentation is also a sufficient basis. The lists in both these experiments comprised two types of items: the crucial items, those which appeared only once per trial for both lists, and the interference items, those whose frequency of appearance varied from once per trial for one list to three times per trial for the second list. When the crucial and interference items both occurred only once per trial, the condition was designated as symmetry presentation and when the interference items occurred more often than the crucial items, the condition was designated as asymmetry. Peformance on both lists was measured only on the crucial items. In the 1965 study, only low similarity lists, made up of trigrams as stimuli and nouns as responses, were used. The major

finding was that when interference items were presented three times as often as the crucial items on each trial, learning of the crucial items was faster than when the interference items occurred only once per trial. In the 1967 experiment, symmetrical vs. asymmetrical presentation of the interference items was compared under both low and high similarity conditions (in this case, each item was similar to all the other items in the list). Under both conditions, performance on the crucial items was superior with asymmetrical presentation of the interfering items than with symmetrical presentation. The facilitation in both these experiments was again restricted to performance following the first correct response and was accompanied by an increase in the proportion of within-crucial-item generalization errors. The type of evidence for grouping in these two experiments is thus the same as in the similarity studies. The basis for grouping, however, is not so clear-cut; either differential frequency of appearance or differential associative strength could be the cue which serves to differentiate the items.

Related Studies

All of the studies just reported were primarily

concerned with interpair grouping as a means of reducing interference within a single list. Several other studies, however, have demonstrated the same type of facilitative effects through rather complex transfer situations. In most of the experiments to be reported, the tested hypothesis was that intratask interference, or more precisely, overcoming intratask interference, is a major source of intertask facilitation (cf. Battig, 1968). As Underwood (1954) has pointed out, the major difficulty in testing this hypothesis is that with variation in stimulus materials, the amount of intertask interference varies concomitantly with intratask interference. Battig has recognized this difficulty and has, therefore, primarily used procedural manipulations, e.g., degree of learning, to vary the amount and sources of interference within a task independent of that between tasks. Since these transfer studies tend to be very complex in both the type of manipulation made and the ways in which the effects are assessed, no attempt will be made to present the experimental conditions or the results in any detail. Instead, only those variations that have resulted in facilitative effects attributed to interpair grouping will be presented.

Grouping on the basis of differential prior practice. When first list pairs were substituted for half of the new pairs in a second paired-associate list, Brown (1964) found that performance on the remaining second-list pairs could either be facilitated or interfered with, depending upon whether degree of prior practice on the substituted pairs was high or low. In a follow-up study (Brown and Battig, 1966) these second-list effects of degree of first-list practice were examined when the first-list-pairs were added to (rather than substituted for) a constant number of second-list-pairs. The results indicated that when pairs added to the second-list had been learned to a substantially higher degree than other first-list-pairs (i.e., easy pairs), _S_s grouped these pairs together on the basis of learning difficulty, and thereby reduced the amount of interpair interference during learning of the second-list.

Grouping on the basis of serial-position constancy. Two studies have reported evidence for grouping on the basis of this cue. Brown and Battig, (1962, Expt. 1) have shown that fewer errors are made in a paired-associate task if pairs which have been given correctly are held constant in the same serial position in which

the first correct response was made, and the serial positions of the incorrect items varied from trial to trial, than under the usual varied-order procedure. In a later experiment (Battig, Brown, and Nelson, 1963, Expt. V), each item was presented in the same serial position until it had been correctly responded to once, after which it was varied in serial position in subsequent trials. Both this "contingent" condition and a "constant" condition, in which items were always presented in the same order, proved to be superior to a condition in which the order for items always varied from trial to trial. This facilitation of the "contingent" groups in both experiments was interpreted as being a result of being better able to group the items on the basis of learned vs. unlearned pairs because of the shift from varied to constant (and constant to varied in the second experiment) serial position, and hence, reduce interpair interference.

Grouping on the directionality of association. In an experiment comparing unidirectional vs. bidirectional conditions of association (this latter condition involves the unsystemmatic trial-to-trial variations in interpair directionality, such that each item serves

both as a stimulus and response, and associations are required in both directions). Schild and Battig (1966) found that if a shift from bidirectional presentations to a constant unidirectional pairing is made contingent upon responding to the pair correctly, performance on the remaining items is facilitated. The cue facilitating S's discrimination between, and grouping of, learned and unlearned pairs, in this case, is the constant directionality of the learned pairs.

Grouping on the basis of added stimulus elements. In an experiment by Brown, Battig, and Pearlstein (1965), a second and third letter were added to an initial single-letter stimulus term for a given pair immediately following attainment of a specified low or high performance criterion (one or three successive correct responses) for that pair. Successive addition of new stimulus letters produced fewer errors to a criterion of three successive errorless trials on the entire list than a group which learned the list with three-letter stimulus terms throughout. This result the investigators interpreted as suggesting that the Ss had categorized the stimuli at various stages of learning on the basis of the number of letters composing

the stimulus.

In the preceding studies, the basis on which grouping of stimulus items occurs has been interpreted as differential learning difficulty or degree of learning. The cues of serial position, bidirectionality, added letter, etc., presumably serve to further differentiate the items into learned and unlearned categories, thus further restricting the sources of interference from the whole list to just members of the same subset. One rather interesting finding from a similar study (Fallon and Battig, 1964) was that grouping on the basis of item difficulty can even override grouping on some other basis, in this case, formal conceptual grouping, as was shown in a post-experimental sorting task by the types of sorts the Ss made. As Battig (1966) has shown, grouping on the basis of degree of learning has some rather important implications for the study of transfer with mixed lists. In the usual mixed design, the amount of transfer under a standard A-B; A-C paradigm (first-list stimuli with new responses) is evaluated in terms of the transfer from A-B to C-D, where C-D are entirely new pairs. In this experiment, Battig varied the number of C-D pairs in the second list and was able to show a shift from "negative" to

"positive" transfer as the number of C-D pairs increased. The facilitation in this case was again attributed to the differentiation of items into "old" and "new" categories and the consequent delimitation of sources of interference. Obviously, then, if this is the case, the amount of negative transfer in this kind of paradigm can either be overestimated or underestimated depending upon the number of control items used.

<u>Summary</u>

In this chapter, most of the studies which have shown evidence for the organizational process, interpair grouping, were reviewed in terms of the types of variables on which the grouping of items supposedly occurs: differential stimulus similarity, differential frequency of appearance, and differential degree of learning.

The facilitative effects of interpair grouping has been attributed to a selective reduction of interference present in the list. More specifically, the <u>S</u> is said to respond to obvious relationships among the stimulus items (differential frequency, for example) as a basis on which to differentiate the items into nonoverlapping subsets, in such a way that the interference in learning

a single item comes primarily from the other items within the same subset. Until the Runquist (in press (b)) study, the type of evidence used most often in support of this notion was the finding that after the first correct response, the highest proportion of overt errors were generalization errors from within a given subset. The Runquist reversal shift study showed that if the "groupings" were maintained but the specific S-R pairings within that group reversed, learning performance was superior to another condition in which S-R pairings were reversed across groups.

Purpose of Present Research

This thesis research represents yet another attempt to obtain some direct evidence for the interpair grouping process in paired-associate learning. The specific problem addressed is again whether the stimulus and response terms are organized into subsets on the basis of "groupable" stimulus-characteristics. However, instead of inferring both the existence and effects of organization through facilitated performance on some second task, an attempt is made to assess the amount of grouping that occurs more directly by including a free recall test following paired-associate learning in the procedure. If the responses are grouped on

the basis of similarity relationships among the stimulus terms these response "groupings" should be manifested as clusters on the free recall tests. The means by which organizational effects in this experiment will be evaluated, then, is through a measure of clustering in free recall.

In this experiment Ss first learn a paired-associate list in which the stimulus terms are divided into different conceptual categories and responses are unrelated adjectives, and then recall either just the stimulus terms or just the response terms on a single free recall test. Clustering of the stimuli and the responses with which they are paired is evaluated in terms of stimulus category membership. Although clustering of stimulus terms might be predicted on a number of bases, clustering of the response terms on the basis of stimulus category membership could only occur through their association with these categories. If the stimulus and response terms are indeed organized on the basis of conceptual category membership, it could be hypothesized that a manipulation which emphasized the existing relationships among the stimulus items should also enhance the amount of clustering of the stimulus- and response-terms. In

this experiment, the relationship between stimulus items is accentuated by several pretraining trials on only the stimulus terms.

Chapter IV

Method

The purpose of this experiment was to determine the presence and amount of organization of stimulus and response terms in paired-associate learning. The Ss learned a paired-associate list having had either pretraining or no pretraining on the stimulus members, and were then tested for recall of either just the stimulus term or just the response term on the first free recall trial, and the other member of the pair on the second free recall trial.

Material

Four 16-item paired-associate lists were used in this experiment. Two were categorized lists, designated as C lists, in which the stimulus terms consisted of four nouns of high taxonomic frequency, as determined by the Cohen, Bousfield, and Whitmarsh (1957) norms, from each of four mutually exclusive conceptual categories. The remaining two were noncategorized lists, designated as NC lists, in which the stimulus terms of each were 16 nouns, also of high taxonomic frequency, but each from a different category. Two sets of response terms, each shared by one categorized and one noncategorized list, were both composed of 16 common adjectives

of from medium to high frequency, as determined by the Thorndike-Lorge (1944) count. Construction of the specific stimulus-response pairings within a given list attempted to eliminate obvious associations between stimuli and responses and among response items. In addition to the paired-associate lists, there were also two sets of pretraining lists, designated as S-S, comprising only the stimulus terms from the two C lists and R-R, comprising only the response terms. All lists used in this experiment are presented in Appendix A.

Design

The design of this experiment separates into three distinct stages: pretraining, paired-associate learning, and free recall of the stimulus or response terms.

Prior to free recall tests, there were three main conditions of training, paired-associate learning of the C list (C), paired-associate learning of the NC list (NC), and paired-associate learning of the C list preceded by pretraining trials on the stimulus terms (S-C).

Following paired-associate learning, half the Ss in each condition were asked to recall only the stimulus terms on the first free recall trial (designated with

Table I

Experimental Design

Group	Pretraining	P.A. Learning	Free-Recall Sequence
C-S	--	C_1	S,R,S,R
	--	C_2	
C-R	--	C_1	R,S,R,S
	--	C_2	
NC-S	--	NC_1	S,R,S,R
	--	NC_2	
NC-R	--	NC_1	R,S,R,S
	--	NC_2	
S-C-S	S_1	C_1	S,R,S,R
	S_2	C_2	
S-C-R	S_1	C_1	R,S,R,S
	S_2	C_2	
S-S	S_1	--	S,S
	S_2	--	
S-R	R_1	--	R,R
	R_2	--	

an -S following the condition label), and the other half only the response terms (designated with -R).

Two additional groups, referred to as S-S and R-R, received a series of pretraining trials on just the stimulus terms from the C lists and just the response terms, respectively, and then were given the free recall test without the interpolation of the paired-associate task.

Two lists (designated with the subscripts 1 and 2) were used under each condition, thus bringing the total number of groups in this experiment to 16 with 10 Ss in each condition. The design of this experiment is illustrated in Table 1.

Subjects

The Ss used in this experiment were 160 students from the Introductory Psychology course, serving in order to obtain course credit. Ss were assigned to one of 16 conditions in the order of their appearance at the laboratory, according to a prearranged running schedule. The running schedule was arranged in ten blocks of 16 conditions, with the order of conditions randomized within each block. An additional five Ss were discarded from the experiment for various reasons: two for failure to comply with instructions, two

because of apparatus failures, and one because of scoring mistakes on the part of the E.

Apparatus

Lists were projected onto an 8 x 8 in. translucent window by means of a 16-mm. Dunning Animatic filmstrip projector. The S sat in a three-sided wooden booth, separating him from both the experimenter and the projector, with the screen being centered about two feet in front of where he sat, approximately at eye level.

Procedure

Pretraining. During the pretraining trials, items (stimulus-or response-term) were exposed one at a time for 2 sec. There was no interval between successive exposures. S was instructed to read each item aloud and try to remember it. After each trial, the projector was stopped, and S told to recall orally as many words as he could. No time limit was given. (All instructions appear in Appendix B) S was then given another opportunity to study the words, and then another test. This series of alternating study and test trials continued until S reached a criterion of 16/16 correct on one trial. There were four orders of presentation (Appendix C) used, the same for all lists.

Selection of these orders attempted to meet the following criteria: (i) no item from one category should be immediately followed by another from the same category; (ii) the last item in one sequence should not appear as the first item presented on the succeeding study trial; (iii) no two items should occur in exactly the same order twice on any of the study trials; and (iv) no item should appear in the same serial position more than once. There were minimum violations of these criteria: one instance in criterion (i), one instance in criterion (iii), and one in criterion (iv). Starting order for each S was determined systematically prior to the experiment.

Paired-associate learning. Immediately after pretraining, the S was read standard paired-associate instructions and his questions about procedure answered. The paired-associate lists were presented by the anticipation method with a 2 sec. exposure of the stimulus term followed by a 2 sec. exposure of the stimulus and response-terms together. There was a 4 sec. blank interval between trials. The four orders of presentation used during pretraining were also used in the paired-associate task. Ss with pretraining started on what would have been the next order had they

required more pretraining trials. Starting order for Ss without pretraining trials was determined by a systemmatic schedule arranged prior to the experiment.

Since S had not yet had an opportunity to see the response that was paired with each stimulus on the first anticipation trial, scoring was not started until the second trial. Anticipation trials were continued until each S had reached a criterion of 14/16 correct responses on a single trial.

Rest interval activity. Immediately following the pretraining trials for the S-S and R-R groups, and after paired-associate learning for all other groups, Ss were given several arithmetic problems to solve as a means to prevent rehearsal of items and to minimize primacy and recency effects on the free recall test. Ss were given a pencil and a sheet of paper which contained a series of 15 arithmetic sequences, and were instructed to try to predict what number should come next in the sequence and then to write that number in a blank space following each sequence. A period of 3-min. was given for the S to complete as many sequences as he could. None were able to finish the task in the allotted time. The data from these arithmetic sequences were not scored.

Free-recall tests. The control groups, S-S and R-R, received two free-recall test trials and all other groups four free-recall trials. S was allowed 1.5 min. for each recall, and recall was oral.

On the first recall, the S-S group was instructed to recall out loud as many of the nouns from the list as they could, starting at a signal from E. Ss were told that there were 16 items in the list. Responses were recorded by the E in the order in which they were emitted, and the time to completion noted. Recall was stopped at the end of 1.5 min. and instructions for the S to recall the same words once again were immediately given. The second recall test started 30 sec. after the termination of the first test and again took 1.5 min. Recall of the adjectives for the R-R groups followed the same procedure.

The recall procedure for the other groups was basically the same. For these groups, however, there were two recalls of both the stimulus terms and the response terms. Half of the conditions (those designated -S) were asked for the nouns on the first recall, the adjectives on the second recall test, the nouns again on the third, and the adjectives again on the last trial. In the other conditions (designated

with -R), Ss first recalled the adjectives, then the nouns, and so on.

Ss were given no feedback on the recall trials.

Chapter V

Results

Pretraining

The mean number of trials taken to reach the criterion of all 16 items correctly given on one trial (TTC) for all groups receiving pretraining is shown in Table II.

TABLE II

MEAN NUMBER OF TRIALS TO CRITERION DURING PRETRAINING

List	S-S	S-C-S	S-C-R	R-R
1	4.2	4.2	3.6	6.9
2	3.2	4.1	2.9	8.5

An analysis of variance (Table IV) over the four similarity groups and the two lists showed a significant difference ($p < .01$) between similarity groups, but no effect due to different lists ($F < 1$), nor any interaction between lists and similarity groups.

Since there was no evidence for a difference between lists, the data from the two lists were combined for the remaining analyses. A Duncan's Multiple Range Test showed that the response-pretraining groups required significantly ($p < .01$) more trials to criterion than any of the stimulus pretraining groups. None of

the latter differed from one another.

Paired-Associate Learning

Trials to criterion. The mean number of trials to reach the criterion of 14/16 correct responses for each group is shown in Table III.

TABLE III

MEAN NUMBER OF TRIALS TO CRITERION DURING PAIRED-ASSOCIATE LEARNING

List	NC-S	NC-R	C-S	C-R	S-C-S	S-C-R
1	9.6	8.1	10.1	10.2	11.2	8.4
2	7.9	10.3	13.6	9.8	9.9	9.5

An analysis of variance (Table IV) performed on these data gave no evidence of true differences between the main treatment groups, between groups differing in which term was recalled first, nor between the two lists used under each condition ($p > .05$ for all differences). Neither was there evidence of an interaction between any of these variables. Even when the list and free recall order data were combined, a Duncan's Multiple Range Test still showed no evidence of any differences between the three main treatment groups. Since the differences between the two lists used under each condition were not significant, the

TABLE IV

SUMMARIES OF ANALYSIS OF VARIANCE

OF TRIALS TO CRITERION

Pretraining

Source	SS	df	MS	F	p
Groups (G)	248.10	3	82.70	16.81	$<.01$
Lists (L)	.05	1	.05	<1	
Groups x Lists (GxL)	20.25	3	6.75	1.37	
Error	354.40	72	4.92		
TOTAL	622.80	79			

Paired-associate Learning

Source	SS	df	MS	F	p
Groups (G)	77.12	2	38.56	1.84	$>.05$
Lists (L)	9.64	1	9.64	<1	--
Term Recalled (T)	30.00	1	30.00	1.44	$>.05$
Groups x Lists (GxL)	15.11	2	7.56	<1	--
Groups x Term (GxT)	31.85	2	15.93	<1	--
Lists x Term (LxT)	4.80	1	4.80	<1	--
Groups x Lists x Term (GxTxL)	85.65	2	42.83	2.05	$<.05$
Error	2254.20	108	20.87		
TOTAL	2508.37	119			

data from these lists were combined for all remaining analyses. Since there were no differences in trials to criterion for groups differing in free recall order, later differences between stimulus and response recall and clustering cannot be attributed to differences in acquisition performance.

Error data. As in earlier studies, the error data that are relevant to the notion of the "grouping" of response terms on the basis of stimulus categories are the percentages of total overt errors that are generalized errors for the groups learning the categorized lists. Since the response terms in the noncategorized list were the same as those in the categorized lists, it was possible to estimate the frequency with which these same errors would be expected to occur if the stimulus terms had not been categorized. For example, the response terms "common", "other", "severe", and "better", were paired respectively with the stimulus terms "John", "Dick", "George", and "Harry" in the categorized list, and with "knife", "golf", "George", and "valley" in the noncategorized list. Confusions within these four pairs with the categorized lists would represent generalized errors that could be a result of response grouping; with the

noncategorized lists, they would represent the frequency with which these same substitutions would be expected on the basis of chance pairings. The proportion of this type of error to the total overt errors made was computed for each S. The mean per cent "generalized" errors for the C, S-C, and NC groups respectively were 59.60%, 63.39%, and 15.98%. An analysis of variance (Table V) on these error data confirmed that the overall difference between these groups was significant ($p < .01$), and a follow-up Duncan's Test showed that while the difference between the C and S-C groups was not significant, both these groups made significantly more "generalized" errors than the NC group.

TABLE V

ANALYSIS OF VARIANCE OF PROPORTION OF GENERALIZED ERRORS

Source	SS	df	MS	F
Between groups	5.553	2	2.777	67.73*
Within groups (error)	4.805	117	.041	
TOTAL	10.36	119		

*difference significant at .01 level of probability

In general, then, the acquisition data in this study show the same interpair grouping phenomena found in previous studies of this sort.

Acquisition performance as assessed by the number of trials taken to reach the criterion did not differ significantly for the three main treatment groups. As usual, the categorized list was slightly, although not significantly, more difficult than the noncategorized list, and the effect of the prior stimulus pretraining was to attenuate this difference even further.

The overt error data also reflect the usual pattern of results. Over 50% of the total overt errors in the C and S-C groups were generalized errors, significantly above a chance frequency provided by the NC group.

<u>Free Recall Data</u>

<u>Number of items recalled</u>. Shown in Table VI are the mean number of stimulus and response terms recalled on the free recall trials for the various -S and -R groups. The recall scores for the pretraining and paired-associate groups were analyzed separately.

Consider first those groups which had paired-associates learning. It will be recalled that on the first and third free recall trials, groups NC-S, C-S, and S-C-S recalled only the stimulus terms, and groups

NC-R, C-R, and S-C-R recalled only the response terms. On the second and fourth trial, recall of the other member of the pair was required so that groups NC-S, C-S, and S-C-S recalled the responses, and NC-R, C-R, and S-C-R recalled the stimuli. Recall was analyzed separately for each trial. An inspection of the means in Table VI reveals a remarkably consistent pattern of results across the different trials: stimulus recall was higher than response recall only for the S-C groups, and the ranking of conditions in terms of number recalled remained constant. These observations were fairly well borne out by the analyses of variance performed on the recall data (Table VII). On trials I-IV, the differences between paired-associate groups were highly significant ($p < .01$). There was no overall difference between stimulus and response recall, except on the fourth trial, but there was a significant interaction between groups and stimulus-or-response recall on all the but third trial. Further comparisons among these groups were made by means of a Duncan's Multiple Range Test on each trial. Considering stimulus recall first, the S-C groups, recalled significantly more than the C groups which in turn recalled more than the NC groups on all four trials. All differences were

TABLE VI

MEAN NUMBER OF ITEMS RECALLED

Pretraining Groups

	S-S	R-R
Free Recall I	15.70	14.20
Free Recall II	15.60	14.10

Paired-associate Groups

Free Recall Trial		NC	C	S-C
I	Stimulus Recall	9.75	12.50	14.40
	Response Recall	10.90	11.90	12.60
II	Stimulus Recall	9.20	11.75	14.10
	Response Recall	9.50	11.85	11.90
III	Stimulus Recall	9.70	11.90	14.65
	Response Recall	9.85	11.85	11.90
IV	Stimulus Recall	9.10	12.30	14.75
	Response Recall	10.05	11.35	12.60

TABLE VII

ANALYSES OF VARIANCE OF RECALL SCORES FOR GROUPS HAVING PAIRED-ASSOCIATE LEARNING

		Free Recall I		Free Recall II		Free Recall III		Free Recall IV	
Source	df	MS	F	MS	F	MS	F	MS	F
Groups (G)	2	101.91	22.70**	138.44	22.55**	164.10	31.74**	168.64	24.56**
Term Recalled (T)	1	5.21	1.16	10.80	1.76	8.00	1.55	15.41	3.16*
Groups x Term (GxT)	2	22.01	4.90*	19.30	3.14*	9.74	1.88	24.43	5.01*
Error	114	4.49	----	6.14	----	5.17	----	4.88	----
TOTAL	119								

* significant at .05 level of probability

**significant at .01 level of probability

significant at the .05 level. Response recall was a little more complex. The differences between the S-C and C groups were not significant on any of the trials, but between the S-C and NC groups, the differences were significant on all trials; and differences between the C and NC groups were significant only on the second and third trial. Comparisons between stimulus and response recall for the three groups showed that stimulus recall was significantly higher than response recall only for the S-C groups, and this difference was maintained across trials.

In general, then, the number of words recalled by those groups learning categorized lists was higher than that recalled by the groups learning the non-categorized lists. The effects of stimulus pretraining was restricted only to stimulus recall.

For groups which received pretraining trials only, it was found that recall of the stimulus terms was significantly higher ($p < .01$) than that of the response terms on both free recall trials (Table VIII). A comparison made between the S-S and the S-C-S groups showed that recall was higher for all groups receiving pretraining trials only than for a group having both the pretraining and the paired-associate learning,

TABLE VIII

SELECTED COMPARISONS OF RECALL SCORES FOR GROUPS HAVING PRETRAINING

	Comparisons	Mean Difference	$\underline{t}$	p(two-tailed)
Free Recall I	S-S vs R-R	1.50	4.34	$<.05$
	S-S vs S-C-S	1.30	2.09	$<.05$
Free Recall II	S-S vs R-R	1.50	3.89	$<.05$

t(38df) = 2.09, $p < .05$. Whether this is a result of the S-S groups having a shorter interval between acquisition and recall tests than the S-C-S groups, or because paired-associate learning interferes somewhat with pretraining retention, cannot be ascertained from this data.

Clustering. Of major interest in this study was the amount of clustering that occurred above chance on each of the free recall trials. For each S, the obtained amount of clustering in his output was determined by summing, over the four categories, the number of times that an item in one category was immediately followed by another item belonging to the same category. Since the stimuli consisted of 4 categories with 4 instances of each category, the maximum number of obtained repetitions was 12. Clustering of the stimuli was defined in terms of their membership within one of the four categories; clustering of the responses in the paired-associate groups was defined in terms of the four response terms paired with the members of a given stimulus category, and in the R-R group by this same response "grouping". Finally, clustering of the noncategorized stimuli was defined by the response "groupings" with which they were paired.

t(1641) = 1.85, p <.05. Whether this is a result of the E-R groups having a shorter interval between registration and recall sessions than the [illegible] groups, or between paired-associate learning instructions associated with generating mediators, cannot be ascertained from this data.

Clustering. Of major interest in this study was the amount of clustering that occurred above chance on each of the free recall trials. For each S, the obtained amount of clustering on the trials was determined by counting, over the four categories, the number of times that an item of one category was immediately followed by another item belonging to the same category. Since the stimuli consisted of 4 categories with 4 instances of each category, the maximum number of clustering repetitions was 12. Clustering of the stimuli was assessed in terms of their membership within one of the four categories; clustering of the responses in the [illegible] groups was defined in terms of the [illegible] response terms paired with the members of a given stimulus category, and in the E-R group by this same response 'grouping'. Finally, clustering of the [illegible] was defined by the response 'groupings' with which they were paired.

In addition to the obtained amount of clustering, an estimate of how much clustering would be expected if the output were random was also determined for each S (Bousfield and Bousfield, 1966). The expected value, the number of repetitions expected on the basis of chance, in a single category of recalled items is $E(r_k) = m_k\,(m_k - 1)/n$ where m_k is the number of words recalled in category k, and n is the total number of words recalled. Summing over all the categories, the expected number of repetitions in the list as a whole is therefore $E(r) = \left[(m_1^2 + \ldots + m_k^2)/n\right] - 1$. (E.g., for a perfectly recalled 16-item list with four categories, $E(r) = 3$.) The expected number of repetitions was subtracted from the obtained number of repetitions for each S to yield a difference score, $D = r - E(r)$. All analyses were performed on these difference scores.

Shown in Table IX are the mean clustering difference scores, the values of t, and the level of significance with a two-tailed test for each of the groups on the four recall trials. Although the absolute differences varied across trials, the pattern of results was constant. All the groups which learned a categorized list, pretraining, and/or paired-associate, showed

TABLE IX

MEAN DIFFERENCE BETWEEN OBTAINED AND EXPECTED CLUSTERING SCORES

FOLLOWING PRETRAINING ONLY

	S-S $\overline{D}$	S-S t	R-R $\overline{D}$	R-R t
Free Recall I	8.40	45.65**	-.77	2.18*
Free Recall II	8.78	21.79**	-1.01	4.59**

FOLLOWING PAIRED-ASSOCIATE LEARNING

Free Recall Trial		NC $\overline{D}$	NC t	C $\overline{D}$	C t	S-C $\overline{D}$	S-C t
I	Stimulus Recall	-.47	1.78	3.95	7.36**	7.43	14.54**
	Response Recall	-.33	1.22	1.01	2.90**	2.45	4.49**
II	Stimulus Recall	-.41	2.41*	3.56	8.84**	6.96	14.26**
	Response Recall	.20	.71	1.66	4.27**	3.85	5.56**
III	Stimulus Recall	-.35	1.86	4.78	11.24**	8.08	18.00**
	Response Recall	-.04	.17	1.75	5.27**	4.25	5.97**
IV	Stimulus Recall	-.23	1.00	4.60	12.92**	7.79	18.38**
	Response Recall	-.37	1.59	2.02	6.25**	5.02	7.24**

*difference significant at .05 level of probability, two-tailed test

**difference significant at .01 level of probability, two-tailed test

clustering of the stimulus terms significantly above chance, $p < .01$, in all cases. Those groups which learned categorized paired-associate lists also showed clustering of the responses above chance, again, at the .01 level of probability. For the R-R and NC groups, on the other hand, there was no evidence of above chance clustering of either the responses or the stimuli on the basis defined by E. The finding that the mean clustering for the R-R group was, in fact, significantly lower than the expected value suggests that Ss were using some other basis, incompatible with that chosen by E, on which to organize the responses.

Since the obtained amount of clustering is somewhat confounded by both the total number of words recalled and the number of words within a category recalled, the comparison between groups were performed on the clustering difference scores. Shown in Table X are the summaries of analyses of variance of these difference values for the paired-associate groups on the four trials. Again, although the exact values differed slightly, the same pattern of results occurred. There was some evidence (reported later) that clustering on the recall trials was not completely independent

TABLE X

ANALYSIS OF VARIANCE OF CLUSTERING DIFFERENCE SCORES

FOLLOWING PAIRED-ASSOCIATE LEARNING

		Free Recall I		Free Recall II		Free Recall III		Free Recall IV	
Source	df	MS	F	MS	F	MS	F	MS	F
Groups (G)	2	280.47	77.69*	303.01	79.74*	405.36	114.83*	450.22	134.80*
Term Recalled (T)	1	207.13	57.74*	64.62	17.01*	142.87	40.47*	100.91	30.01*
Groups x Term (GxT)	2	68.84	19.07*	36.02	9.48*	48.22	13.66*	21.47	6.43*
Error	114	3.61	-----	3.80	----	3.53	-----	3.34	----
TOTAL	119	560.05							

*difference significant at .01 level of probability

of prior recall. The results of the analyses of variance showed a significant difference ($p < .01$) in the amount of overall clustering between the three main groups, a significant difference ($p < .01$) between the amounts of clustering of the stimulus and response terms, and a significant interaction ($p < .01$) between these two variables. Independent comparisons between the amounts of stimulus and response clustering were made by t-tests for each of the main treatment groups in order to determine the source of this interaction (Table XI). It was found that stimulus clustering was superior ($p < .01$) to response clustering for both the C and S-C groups, but did not differ significantly from response clustering in the NC group. Two additional comparisons were made by means of a Duncan's Test to determine the differences between groups in the amount of clustering of the stimulus and response terms. Regardless of which term was recalled, clustering was the highest for the S-C groups, second for the C groups and lowest for the NC groups ($p < .05$ for all differences).

The overall results of analyses of variance, t-tests, and Duncan's tests on the remaining free

TABLE XI

COMPARISONS OF CLUSTERING DIFFERENCE MEANS BETWEEN

STIMULUS-AND RESPONSE-TERM RECALL

	Free Recall I		Free Recall II		Free Recall III		Free Recall IV	
Groups	$\overline{D}$	t	$\overline{D}$	t	$\overline{D}$	t	$\overline{D}$	t
NC	- .14	<1	- .61	<1	- .31	<1	.14	<1
C	2.94	4.89*	1.90	3.08*	3.03	5.10*	2.58	4.46*
S-C	4.98	8.29*	3.11	5.04*	3.83	6.45*	2.77	4.79*

*difference significant at .01 level of probability (two-tail test)

recall trials were identical to those on the first trial, and hence require no additional discussion.

As mentioned earlier, there was evidence that the amount of clustering was not independent of prior recall. More specifically, it was found that the amount of response clustering was higher when the stimulus terms had already been recalled. For the groups that recalled the responses on the first free recall trial, i.e., the C-R, S-C-R, and NC-R groups, the mean response clustering scores were 1.01, 2.85, and -.33 respectively. For the groups that recalled the responses on the second free recall trial, i.e., the C-S, S-C-S, and NC-S groups, the mean response clustering scores were 1.66, 3.85, and .20 respectively. The analysis of variance performed on these data (Table XII) showed a significant effect of prior recall ($p < .01$), but no interaction with treatments. Although the response clustering was higher after prior recall for all three groups, only in the S-C groups was the difference significant at the .05 level (Table XIII). While prior recall of the stimulus terms enhanced response clustering, it did not increase the number of response terms recalled. In fact, recall was generally poorer, although not significantly so ($p > .05$, Table XII)

TABLE XII

ANALYSES OF VARIANCE OF RESPONSE RECALL AND RESPONSE CLUSTERING FOR GROUPS HAVING PRIOR VS. NO PRIOR RECALL OF STIMULUS TERMS

Source	df	Recall MS	Recall F	Clustering MS	Clustering F
Prior Recall (P)	1	15.41	2.81	23.87	6.09*
Groups (G)	2	47.66	8.68*	100.46	25.63**
Prior Recall x Groups (PxG)	2	4.56	<1	2.75	<1
Error	114	5.49		3.92	
TOTAL	119				

* difference significant at .05 level of probability

**difference significant at .01 level of probability

TABLE XIII

SELECTED COMPARISONS OF CLUSTERING DIFFERENCE SCORES FOR GROUPS WITH AND WITHOUT PRIOR S RECALL

Comparisons	Mean Difference	t	p(two-tailed)
NC-S vs NC-R	.13	<1	----
C-S vs C-R	.63	1.01	>.05
S-C-S vs S-C-R	1.50	2.40	<.05

when the stimulus terms had been recalled first.

In general, the results showed clustering above chance for both the stimulus- and response-terms following paired-associate learning of a categorized list, and the effect of stimulus pretraining was to increase the amount of clustering with stimuli and responses. There was no evidence of any clustering following learning of a non-categorized list.

The preceding description was concerned with the amount of clustering that occurred on the individual recall trials. Another measure, that of intertrial repetitions, or ITR, (Bousfield and Bousfield, 1966), was used to assess constancies in the order of output over successive recall trials for the various groups. The obtained amount of ITR is acquired by counting, for each S, the number of pairs of items that are recalled in the same order on both trials. This value is compared with the amount of repetition that would be expected if the order of output were random, E(ITR). The expected value is determined by the formula $E(ITR) = C(C - 1)/hk$ where h = the number of items recalled on trial t, k = the number of items recalled on trial t + 1, and C = the number of items common to the two recalls. The difference between

those two values is evaluated by a one-tailed $\underline{t}$-test in order to determine if sequential stability across trials is significantly above chance.

Table XIV shows the mean difference scores, $\overline{D} = \overline{O}(ITR) - \overline{E}(ITR)$, the values of $\underline{t}$, and the level of probability for each of the groups when repetitions are measured between the two stimulus recall trials, the two response recalls, the first stimulus and first response recall, and the second stimulus and response recall. Consider first the amount of repetition that occurred between the two stimulus recall trials. When the stimuli were recalled on the first and third trials, the number of repetitions was significantly above chance for all groups learning the categorized lists, but not for the group learning the noncategorized list. When stimulus recall was preceded by response recall, there was a significant number of repetitions above chance even in the NC-R group.

Similarly, the number of repetitions between response recall trials was significantly above chance, except for the NC-S group, in which response recall was preceded by stimulus recall.

Significantly above chance ITR on the part of the groups learning categorized lists suggests that the

TABLE XIV

MEAN DIFFERENCE BETWEEN OBTAINED AND

EXPECTED INTERTRIAL REPETITION

BETWEEN STIMULUS RECALL TRIALS

	NC		C		S-C		S-S	
Stimuli Recalled	$\bar{D}$	t	$\bar{D}$	t	$\bar{D}$	t	$\bar{D}$	t
Trial I + III	-.06	-.33	1.28	2.94*	3.09	6.79*	3.58	8.80*
Trial II + IV	.58	2.77*	1.49	4.23*	2.59	6.31*		

BETWEEN RESPONSE RECALL TRIALS

	NC		C		S-C		R-R	
Responses Recalled	$\bar{D}$	t	$\bar{D}$	t	$\bar{D}$	t	$\bar{D}$	t
Trial I + III	.66	2.06*	.61	3.32*	.90	2.14*	2.98	4.45*
Trial II + IV	.33	1.19	1.01	3.64*	1.44	4.85*		

BETWEEN FIRST STIMULUS AND FIRST RESPONSE RECALL

	NC		C		S-C	
Term Recalled First	$\bar{D}$	t	$\bar{D}$	t	$\bar{D}$	t
Stimulus	-.11	.70	.37	2.15*	.97	2.52*
Response	.08	.45	.09	.54	.62	1.73*

BETWEEN SECOND STIMULUS AND SECOND RESPONSE RECALL

	NC		C		S-C	
Term Recalled First	$\bar{D}$	t	$\bar{D}$	t	$\bar{D}$	t
Stimulus	.11	.74	.24	1.07	1.53	4.13*
Response	.25	1.12	.39	1.31	.77	2.18*

*difference significant at .05 level of probability with one tailed-test

"organization" of the stimulus and the response terms was fairly stable throughout recall trials. Significant ITR of response recall for the NC and R-R groups suggests that the Ss did organize the responses in some way, even though it was not the type of organization measured by E.

When sequential constancies in order of output were measured between successive stimulus and response recall trials, the number of repetitions was significant for the C-S group only on the first two trials, and for the S-C-S and S-C-R groups, on both sets of trials. This result reflects fairly well the earlier finding that stimulus pretraining increases the amount of response clustering on the basis of the stimulus categories.

One of the major problems with this ITR measure is that it measures only pairwise constancies, and only those in one direction, so that it does not take into account total or partial reversals in the order of output. Thus, if, for example, items "football", "lion", and "table" are recalled in the order 1-2-3 on the first trial, and 3-2-1, 2-1-3, or 1-3-2, etc., on the second trial, these would not be counted as repetitions. In this case, then, the amount of

organization as measured by sequential constancies is probably underestimated.

The final measure of recall performance to be considered is the mean latency of word recall on the first free recall trial (Table XV). For each S, the mean latency per word recalled was computed by dividing the total number of seconds it took him to recall the words by the number of words recalled.

TABLE XV

MEAN LATENCY (IN SECONDS) OF WORD RECALL

Word Recalled	NC	C	S-C	Pretraining
Stimulus	5.40	3.52	2.61	1.61
Response	4.48	4.20	4.68	2.86

An analysis of variance of the latency data (Table XVI) for the groups which had paired-associate learning showed a significant difference between the three treatment groups ($p < .01$), and a significant interaction between groups and the terms recalled ($p < .01$). Individual comparisons between mean stimulus and response word latencies showed that for the NC groups, stimulus word latency was significantly greater than response latency ($p < .01$), but for the C and S-C groups mean

TABLE XVI

SUMMARIES OF ANALYSIS OF VARIANCE AND SELECTED *t*-TESTS ON MEAN WORD LATENCY SCORES FOR GROUPS HAVING PAIRED-ASSOCIATE LEARNING

ANALYSIS OF VARIANCE

Source	SS	df	MS	F
Groups (G)	36.44	2	18.22	6.26*
Term Recalled (T)	11.10	1	11.10	3.81
Groups x Term (GxT)	46.53	2	23.27	8.00*
Error	331.32	114	2.91	
TOTAL	425.39	119		

*difference significant at the .01 level of probability

t-TESTS

Comparison	$\bar{D}$	*t*	p(two-tailed)
NC-S vs NC-R	.92	3.16	$<.01$
C-S vs C-R	-.68	2.34	$<.05$
S-C-S vs S-C-R	-2.07	5.39	$<.01$

mean stimulus latency was significantly less than response latency.

Chapter VI

Discussion

The major findings of this experiment were:

1. Groups that learned categorized paired-associate lists did not differ in learning rate from groups that learned noncategorized lists, despite the fact that conceptual similarity among the stimulus terms in the categorized lists should have produced greater interference.

2. Of the total overt errors made by the groups learning categorized lists, over fifty per cent of these were generalized or confusion errors from within the categories.

3. Groups that learned categorized lists showed evidence of organizing the stimulus terms into subsets on the basis of stimulus category membership, as measured by clustering in free recall.

4. Groups that learned categorized lists also showed evidence of organizing the response terms on the basis of the stimulus category to which its stimulus pair belonged.

5. Organization of the stimulus terms was greater than that of the response terms.

6. Recall was higher for groups that learned a

categorized list than for groups that learned a noncategorized list.

7. Both the amount of organization and the number recalled for the stimulus terms were increased by stimulus pretraining.

8. Organization, but not recall, of the response terms was increased by stimulus pretraining, and this effect was enhanced when stimulus recall preceded response recall.

9. The organization of the stimulus and response terms was stable on a second free recall trial.

Interpair grouping has been conceptualized as a process of subcategorizing a paired-associate list into groups or subsets of pairs on the basis of some common property or interrelationships among the pairs. The effect of this grouping is to restrict the possible sources of interference in learning a single pair to other pairs from within the same subset. In the present experiment, the stimulus terms consisted of several instances of different conceptual categories, and the response terms were unrelated. The question was whether the stimulus and response terms would be organized into subsets on the basis of the conceptual similarity relationships among the stimuli, thereby

limiting the number of responses given to a particular stimulus to those paired with the similar stimuli, and hence, reduce the amount of interference in learning the list.

The existence of the grouping process has usually been inferred from two types of findings: First, generalized or confusion errors from within category have accounted for a much greater proportion of the total overt errors than would be expected on the basis of random responding. Secondly, with lists of this type, acquistion performance has sometimes shown facilitation when this would not normally be predicted. In this experiment, not only was there a high proportion of generalized errors, but the categorized lists, which should have produced greater interference, did not take significantly longer to learn than the noncategorized lists. The error data suggest, then, that most of the interference did come from within a subset, and the slight facilitation in learning the categorized lists, suggests that this restriction of interference was to increase the rate of learning somewhat.

Following paired-associate learning, Ss were given a series of free recall trials on which they were asked

to recall the stimulus terms and the response terms on alternate trials. The rationale was that, if S did group the pairs into subsets during paired-associate learning, these groupings would be reflected in the order in which he recalls the words, i.e., as clustering. This is exactly what happened. Both the stimulus and the response terms in the categorized lists showed clustering significantly above chance when evaluated in terms of the stimulus categories.

The fact that the stimulus terms were clustered in the free recall trials is not surprising. This finding could, in fact, be predicted, not only on the basis of grouping the stimuli into subsets by means of some perceived relationships among them, but by mediated stimulus generalization between items that are conceptually similar. Similarly, the effects of stimulus pretraining on stimulus clustering and recall could have been simply to increase the amount of generalization between the similar stimuli.

There is a similar difficulty in interpreting response term clustering. If S were aware of the categorized nature of the stimuli, clustering of the response terms on the free recall tests could have occurred in two ways. First, the response terms could

have been grouped during paired-associate learning by virture of their associations with members of a particular stimulus category. The alternative interpretation is that S, when asked to recall the responses, first recalled the stimulus term, then said the response term that had been paired with it out loud. With this latter interpretation, clustering of the responses would be observed to the extent that the similar stimuli were recalled together. What is envisioned here is that the recall of one stimulus (S_{A1}) from a given category (A) leads to the recall of both the response term that was paired with it (R_1), and another stimulus term (S_{A2}) from the same category. Similarly, S_{A2} leads to the recall of R_2 and S_{A3}, and so on. Although the data do not allow a clear choice between these two interpretations, there are a few results that are difficult to account for if the second alternative, stimulus-cued responding, was the only way in which the responses were recalled. Consider first the fact that, while there was no evidence of a difference between the number of stimuli and responses recalled by the groups learning the categorized lists, clustering of the stimulus terms was significantly greater than that of the response terms. If the

stimulus terms were used as cues for recalling the responses, one could predict that the amount of response clustering should be almost as great as the stimulus clustering. There is an argument against this criticism, however, and that is, that the recall of the response term interrupts the generalizing between similar stimulus terms so that a stimulus from another category is recalled next. A second piece of evidence that could be used against the notion of stimulus-cued responding comes from ITR data between consecutive stimulus and response recall trials. To the extent that response recall is dependent upon stimulus recall, one could expect the order of output for each to be somewhat similar. The only groups that showed a significant number of intertrial repetitions between stimulus and response recalls consistently, however, were the categorized list plus pretraining conditions, despite the fact that the groups learning only the categorized paired-associate lists showed significant ITR both between the two stimulus recalls and the two response recalls. A final piece of evidence comes from the latency data (Table XV). For the NC groups, it took significantly longer to recall a stimulus term than to recall a response term. For this

group at least, then, it seems unlikely that the stimulus terms were used as cues for response recall. If one makes the somewhat tenuous assumption that the latency of word recall for the stimulus terms would be the same whether recalled orally (i.e., on the stimulus recall trials) or subvocally (i.e., as a cue for response recall), then, with the notion of stimulus-cued response recall, the average length of time to output the response term would be only .68 seconds for the C group, and 2.07 seconds for the S-C group. These figures are obtained by subtracting the mean word latency for stimulus recall from that for response recall. The mean word latency for the R-R group, having had a lot of prior practice in outputting the response terms during pretraining, was 2.86 seconds. If the latency for the R-R group is taken as an estimate of the least amount of time required to output a response term, it can be seen that stimulus-cuing would take up too much time.

There are obviously two major problems with all these criticisms. First, they could only be valid if it were assumed that stimulus-cued response recall was the only thing that was happening during response recall and it is doubtful that anyone would hold to

such an extreme position. Secondly, it is doubtful that all the assumptions underlying these criticisms are justified. Furthermore, even if these criticisms are warranted, there is no evidence that convincingly demonstrates the first interpretation, i.e., response clustering on the basis of response groupings formed during paired-associate learning, accounts for all the recall data either.

Two other findings should also be briefly discussed. The first of these was that the amount of clustering for the R-R group, when assessed in terms of the response subsets that should develop during the learning of the categorized list, was significantly lower than that expected if output had been random. This finding could be interpreted as showing that the S did organize the response terms, but in a way that was incompatible with the type of organization measured by E. An inspection of the recall protocols supports this interpretation. Most of the Ss tended to recall the words alphabetically, at least to the extent of recalling words that started with the same letter together. The responses within the E-defined categories, on the other hand, all started with different letters. (There was also some evidence of this type of

organization following the learning of the categorized list; as a result, the amount of response grouping following paired-associate learning may have been underestimated.) The significant ITR for this group reflects the stability of this alphabetical organization over the two recall trials. A bit more difficult to interpret is the significant ITR between the two response recall trials for the NC group that occurred when the response terms were recalled first, but not when the stimulus terms were recalled first. It is possible that the responses had been organized on some basis during paired-associate learning but there had been no opportunity for this organization to be manifested until the free recall trial. The other possibility was that the responses were not organized until the free recall stage. The only evidence that supports the first interpretation is that there were some confusion errors between response terms starting with the same letter during paired-associate learning, and these response terms tended to be recalled together during the free recall trials. These errors, however, did not account for a very large proportion of the total errors, and responses starting with the same letters that were not substituted for each other during

learning were also frequently recalled together on the free recall trials.

Before concluding, one limitation of this experiment should be pointed out. There is little doubt that clustering on the free recall trials reflects grouping information that was acquired during paired-associate learning. It cannot be inferred, however, from the free recall data, that this information had any effect on paired-associate performance.

Insofar as the clustering in free recall technique is able to assess the effects of organization occuring in paired-associate learning, this experiment does provide some further evidence for the interpair grouping process. More specifically, evidence was found for the categorization of stimulus and response terms into subsets on the basis of similarity relationships among the stimuli.

References

Battig, W. F. A shift from "negative" to "positive" transfer under the A-C paradigm with increased number of C-D control pairs in a mixed list. Psychon. Sci., 1966, 4, 421-422.

Battig, W. F. Paired-associate learning. In T. R. Dixon & D. L. Horton (Eds.) Verbal behavior and general behavior theory. Englewood Cliffs: Prentice-Hall, 1968.

Battig, W. F., Brown, S. C., & Nelson, D. Constant vs. varied serial order in paired-associate learning. Psychol. Rep., 1963, 12, 695-721.

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. J. gen. Psychol., 1953, 49, 229-240.

Bousfield, W. A., Berkowitz, H., & Whitmarsh, G. A. Associative clustering in the recall of minimally meaningful geometric designs. Canad. J. Psychol., 1959, 13, 281-287.

Bousfield, A. K., & Bousfield, W. A. Measurement of clustering and of sequential constancies in repeated free recall. Psychol. Rep., 1966, 19, 935-941.

Bousfield, W. A., & Cohen, B. H. The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. J. Psychol., 1953, 36, 67-81.

Bousfield, W. A., & Cohen, B. H. Clustering in recall as a function of the number of word-categories in stimulus-word lists. J. gen. Psychol., 1956, 54, 95-106.

Bousfield, W. A., Cohen, B. H., & Whitmarsh, G. A. Associative clustering in the recall of words of different taxonomic frequencies of occurrence. Psychol. Rep., 1958, 4, 39-44.

References

Besch, N. F. A shift from "negative" to "positive" transfer under the A-C paradigm with increased number of [illegible] in a mixed list. Psychon. Sci., 1966, 6, 331-332.

Battig, W. F. Paired-associate learning. In T. R. Dixon & D. L. Horton (Eds.) Verbal behavior and general behavior theory. Englewood Cliffs: Prentice-Hall, 1968.

Battig, W. F., Brown, S. C., & Nelson, D. Constant vs. varied serial order in paired-associate learning. Psychol. Rep., 1963, 12, 695-721.

Bousfield, W. A. The occurrence of clustering in the recall of randomly arranged associates. J. gen. Psychol., 1953, 49, 229-240.

Bousfield, W. A., Berkowitz, H., & Whitmarsh, G. A. Associative clustering in the recall of minimally meaningful geometric designs. Canad. J. Psychol., 1959, 13, 281-287.

Bousfield, A. K., & Bousfield, W. A. Measurement of clustering and of sequential constancies in repeated free recall. Psychol. Rep., 1966, 19, 935-942.

Bousfield, W. A., & Cohen, B. H. The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. J. Psychol., 1953, 36, 67-81.

Bousfield, W. A., & Cohen, B. H. Clustering in recall as a function of the number of word-categories in stimulus-word lists. J. gen. Psychol., 1956, 54, 95-106.

Bousfield, W. A., Cohen, B. H., & Whitmarsh, G. A. Associative clustering in the recall of words of different taxonomic frequencies of occurrence. Psychol. Rep., 1958, 4, 39-44.

Bousfield, W. A., & Puff, C. R. Clustering as a function of response dominance. J. exp. Psychol., 1964, 67, 76-79.

Bousfield, W. A., & Puff, C. R. Determinants of the clustering of taxonomically and associatively related word pairs. J. gen. Psychol., 1965, 73, 211-221.

Bousfield, W. A., & Sedgewick, C. H. W. An analysis of sequences of restricted associative responses. J. gen. Psychol., 1944, 30, 149-165.

Bousfield, W. A., Steward, J. R., & Cowan, T. M. The use of free associational norms for the prediction of clustering. J. gen. Psychol., 1964, 70, 205-214.

Bousfield, W. A., Whitmarsh, G. A., & Berkowitz, H. Partial response identities in associative clustering. J. gen. Psychol., 1960, 63, 233-238.

Brand, H., & Woods, P. J. The organization of the retention of verbal material. J. gen. Psychol., 1958, 58, 55-68.

Brown, S. C. Interpair interference as a function of level of practice in paired-associate learning. J. exp. Psychol., 1964, 67, 316-323.

Brown, S. C., & Battig, W. F. Partial serial-position constancy in paired-associate lea rning. J. verb. Learn. verb. Behav., 1962, 1, 42-47.

Brown, S. C., & Battig, W. F. Second-list paired-associate facilitation produced by addition of previsouly learned first-list pairs. J. verb. Learn. verb. Behav., 1966, 5, 320-321.

Brown, S. C., Battig, W. F., & Pearlstein, R. Effect of successive addition of stimulus elements on paired-associate learning. J. exp. Psychol., 1965, 70, 87-93.

Cofer, C. N. A study of clustering in free recall based on synonyms. J. gen. Psychol., 1959, 60, 3-10.

Cofer, C. N. On some factors in the organizational characteristics of free recall. Amer. Psychologist, 1965, 20, 261-272.

Cofer, C. N. Some evidence for coding processes derived from clustering in free recall. J. verb. Learn. verb. Behav., 1966, 5, 188-192.

Cofer, C. N., & Bruce, D. R. Form-class as the basis for clustering in the recall of nonassociated words. J. verb. Learn. verb. Behav., 1965, 4, 386-389.

Cofer, C. N., Bruce, D. R., & Reicher, G. M. Clustering in free recall as a function of certain methodological variations. J. exp. Psychol., 1966, 71, 858-866.

Cohen, B. H. An investigation of recoding in free recall. J. exp. Psychol., 1963, 65, 368-376. (a)

Cohen, B. H. Recall of categorized word lists. J. exp. Psychol., 1963, 66, 227-234. (b)

Cohen, B. H. Some-or-none characteristics of coding behavior. J. verb. Learn. verb. Behav., 1966, 5, 182-187.

Cohen, B. H., Bousfield, W. A., & Whitmarsh, G. A. Cultural norms for verbal items in 43 categories. Technical Report No. 22, 1957, University of Connecticut, Contract Nonr - 631(00), Office of Naval Research.

Cramer, P. Mediated clustering and importation with implicit verbal chains. Psychon. Sci., 1965, 2, 165-166.

Dallett, K. M. Number of categories and category information in free recall. J. exp. Psychol., 1964, 66, 1-12.

Fallon, D., & Battig, W. F. Role of difficulty in rote and concept learning. J. exp. Psychol., 1964, 68, 85-88.

Gonzales, R. C., & Cofer, C. N. Exploratory studies of verbal context by means of clustering in free recall. J. genet. Psychol., 1959, 95, 293-320.

Jenkins, J. J., Mink, W. D., & Russell, W. A. Associative clustering as a function of verbal association strength. Psychol. Rep., 1958, 4, 127-136.

Jenkins, J. J., & Russell, W. A. Associative clustering during recall. J. abn. soc. Psychol., 1952, 47, 818-821.

Joinson, P. A., & Runquist, W. N. Similarity, degree of learning and retention. J. verb. Learn. verb. Behav., in press.

Kendler, H. H. Coding: associationalistic or organizational? J. verb. Learn. verb. Behav., 1966, 5, 198-200.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), The psychology of learning and motivation. New York: Academic Press, 1967.

Mandler, G., & Pearlstone, Z. Free and constrained concept learning and subsequent recall. J. verb. Learn. verb. Behav., 1966, 5, 126-131.

Marshall, G. R. The organization of verbal material in free recall: the effects of patterns of associative overlap on clustering. Unpublished doctoral dissertation, NYU, 1963. Referred to in: Cofer, C. N. On some factors in the organizational characteristics of free recall. Amer. Psychologist, 1966, 20, 261-272.

Marshall, G. R., & Cofer, C. N. Associative, category, and set factors among word pairs and triads. Technical Report No. 4, 1961, NYU. Referred to in: Cofer, C. N. On some factors in the organizational characteristics of free recall. Amer. Psychologist, 1965, 20, 261-272.

Gonzalez, R. C., & Cofer, C. N. Exploratory studies of verbal context by means of clustering in free recall. J. genet. Psychol., 1959, 95, 293-320.

Jenkins, J. J., Mink, W. D., & Russell, W. A. Associative clustering as a function of verbal association strength. Psychol. Rep., 1958, 4, 127-136.

Jenkins, J. J., & Russell, W. A. Associative clustering during recall. J. abn. soc. Psychol., 1952, 47, 818-821.

Johnson, R. A., & [illegible], W. W. [illegible], forms of learning and retention. J. verb. Learn. verb. Behav., in press.

Mandler, G. [illegible] J. verb. Learn. verb. Behav., 1966, 5, 198-200.

Mandler, G. Organization and memory. In K. W. Spence & J. T. Spence (Eds.), The psychology of learning and motivation. New York: Academic Press, 1967.

Mandler, G., & Pearlstone, Z. Free and constrained concept learning and subsequent recall. J. verb. Learn. verb. Behav., 1966, 5, 126-131.

Marshall, G. R. The organization of verbal material in free recall: The effects of patterns of associative overlap on clustering. Unpublished doctoral dissertation, NYU, 1963. Referred to in: Cofer, C. N. On some factors in the organizational characteristics of free recall. Amer. Psychologist, 1965, 20, 261-272.

Marshall, G. R., & Cofer, C. N. Associative, category, and set factors in clustering among word pairs and triads. Technical Report No. 4, 1961, NYU. Referred to in: Cofer, C. N. On some factors in the organizational characteristics of free recall. Amer. Psychologist, 1965, 20, 261-272.

Mathews, R. Recall as a function of number of classificatory categories. J. exp. Psychol., 1954, 47, 241-247.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity to process information. Psychol. Rev., 1956, 63, 81-97.

Postman, L. Hermann Ebbinghaus. Amer. Psychologist, 1968, 23, 149-157.

Rosenberg, S. Associative clustering and repeated trials. J. gen. Psychol., 1966, 74, 89-96.

Runquist, W. N. Order of presentation and number of items as factors in paired-associate verbal learning. J. verb Learn. verb. Behav., 1965, 4, 535-540.

Runquist, W. N. Intralist interference as a function of list length and interstimulus similarity. J. verb. Learn. verb. Behav., 1966, 5, 7-13.

Runquist, W. N. Effects of variable frequency of presentation on paired-associate learning as a function of interstimulus similarity. J. verb. Learn. verb. Behav., 1967, 6, 470-475.

Runquist, W. N. Formal intralist stimulus similarity in paired-associate learning. J. exp. Psychol., in press. (a)

Runquist, W. N. Reversal vs. non-reversal repairing in categorized paired-associates lists. J. verb. Learn. verb. Behav., in press. (b)

Runquist, W. N., & Joinson, P. A. Rated similarity of low association value trigrams. J. verb. Learn. verb. Behav., in press.

Schild, M. E. ,& Battig, W. F. Directionality in paired-associate learning. J. verb. Learn. verb. Behav., 1966, 5, 42-49.

Thorndike, E. L., & Lorge, I. The Teacher's word book of 30,000 words. New York: Columbia University Press, 1944.

Tulving, E. Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.), Verbal behavior and general behavior theory. Englewood Cliffs: Prentice-Hall Inc., 1968.

Underwood, B. J. Intralist similarity in verbal learning and retention. Psychol. Rev., 1954, 61, 160-166.

Underwood, B. J., Ekstrand, B. R., & Keppel, G. An analysis of intralist similarity in verbal learning with experiments on conceptual similarity. J. verb. Learn. verb. Behav., 1965, 4, 447-462.

Wicklund, D. A., Palermo, D. S., & Jenkins, J. J. Associative clustering in the recall of children as a function of verbal association strength. J. exp. child. Psychol., 1965, 2, 58-66.

APPENDICES

APPENDIX A

Learning Materials

	Stimulus words			Response words
	C list	NC list		both lists
List 1				
1	football	truck	-	civil
2	hockey	piano	-	sandy
3	tennis	ruby	-	northern
4	golf	golf	-	thankful
5	dog	hour	-	actual
6	cow	cow	-	extreme
7	bear	lamp	-	willing
8	lion	bean	-	frequent
9	wine	rain	-	tender
10	gin	knife	-	annual
11	scotch	window	-	noble
12	rye	rye	-	certain
13	shirt	valley	-	easy
14	hat	hat	-	guilty
15	tie	George	-	complete
16	coat	leg	-	unknown
List 2				
1	sofa	cow	-	latter
2	lamp	lamp	-	active
3	table	ruby	-	southern
4	chair	piano	-	total
5	pea	rye	-	upper
6	carrot	window	-	absent
7	bean	bean	-	grateful
8	lettuce	hat	-	maral
9	John	knife	-	common
10	Dick	golf	-	other
11	George	George	-	severe
12	Harry	valley	-	better
13	car	leg	-	simple
14	truck	truck	-	able
15	boat	hour	-	distinct
16	plane	rain	-	blank

APPENDIX B

Instructions Given to Subjects

Pretraining Instructions

"In the first part of the experiment, you will see a series of 16 common nouns (adjectives) presented one at a time on the screen in front of you. When each word appears, you are to read it out loud and try to remember it. We will go through the whole list once with you reading out each word. Don't try to learn the words in order. Just try to learn which words are in the list. After we have gone through the list once, I will test you to see how many words you remember by asking you to say out loud all the words in the list. Then you will have another opportunity to study the words, then another test. We will continue in this way, first a study trial, then a test trial, until you can remember the entire list.

Do you have any questions? Remember to pronounce the words on the study trials. Are you ready to being?"

Paired-associate Instructions

"In this part of the experiment, you will be required to learn to associate some common nouns with some common adjectives.

The list of words you will learn consists of 16 pairs like the pair on this card (_E_ shows _S_ the example card). These pairs will be presented on the screen in front of you. When we begin, the noun will always appear alone for a brief interval on the left-hand side of the screen. Then the adjective will appear next to it on the right-hand side. Your task is to associate or connect the two words so that you will be able to say the adjective _aloud_ while the noun is in the window alone; that is, say the adjective that goes with the noun before the adjective appears.

Since the order in which the pairs follow each other will not always be the same, you must learn the pairs _as pairs_, and not in the order in which they follow each other.

When we start, we will go through the whole list once, so that you can study the pairs and try to make associations between the members of the pairs. After we have gone through the list once, a blank space will appear in the window. The appearance of this blank means that we are starting another trial, in this case, the second trial. It is on the second trial that when the noun appears, you must begin trying

to say the adjective that goes with it <u>before</u> that adjective appears. We will then continue through the list without interruption while you attempt to anticipate the adjectives before they appear. You will continue through the list, trial after trial, until I stop you.

Always try to get as many pairs correct on each trial as you can. If you are having trouble anticipating some of the words or are giving some incorrectly, don't let this discourage you or prevent you from doing the best you can. The lists are long and some are quite difficult, so don't worry about making mistakes.

Are there any questions? All right. Remember to anticipate the adjective just after the noun has appeared. Start as soon as you think you know any of the pairs."

<u>Arithmetic Sequences Instructions</u>

"Now I have something different for you to do. On the paper I am about to give you, there are several sequences of numbers. Your task is to figure out what number should come next in the sequence, and then to write that number in the blank space following the sequence. You will be allowed only a short amount of time for these problems, so please work as quickly

but as accurately as you can.

Any questions? All right, begin."

Free Recall Instructions

"Now I want to see how well you can remember the words that were in the list you learned just a few moments ago.

When I say READY, I want you to say out loud all the nouns (adjectives) that were in the list (that is, the words that appeared on the left- (right-) hand side of the screen.) There were 16 of these. Any questions? READY.

Now when I say READY, I want you to say all the adjectives (nouns) that were in the list, that is, the words that were on the right- (left-) hand side. Any questions? READY.

Now I want you to say all the nouns (adjectives) again. READY.

Finally, I want you to say all the adjectives (nouns) again. READY.

APPENDIX C

Orders of Item Presentation

Order 1	Order 2	Order 3	Order 4
1	16	14	12
13	6	11	6
9	3	4	14
5	10	7	3
6	1	3	16
2	5	12	7
14	9	8	11
10	4	15	8
15	8	5	4
11	14	1	9
3	12	16	13
7	13	9	5
16	11	10	2
8	15	2	15
4	2	6	1
12	7	13	10